Support vector classifier.

In support vector classifier we create Marginal plane along both side of the best fit line in such a way that distance of the marginal plane is maximum.

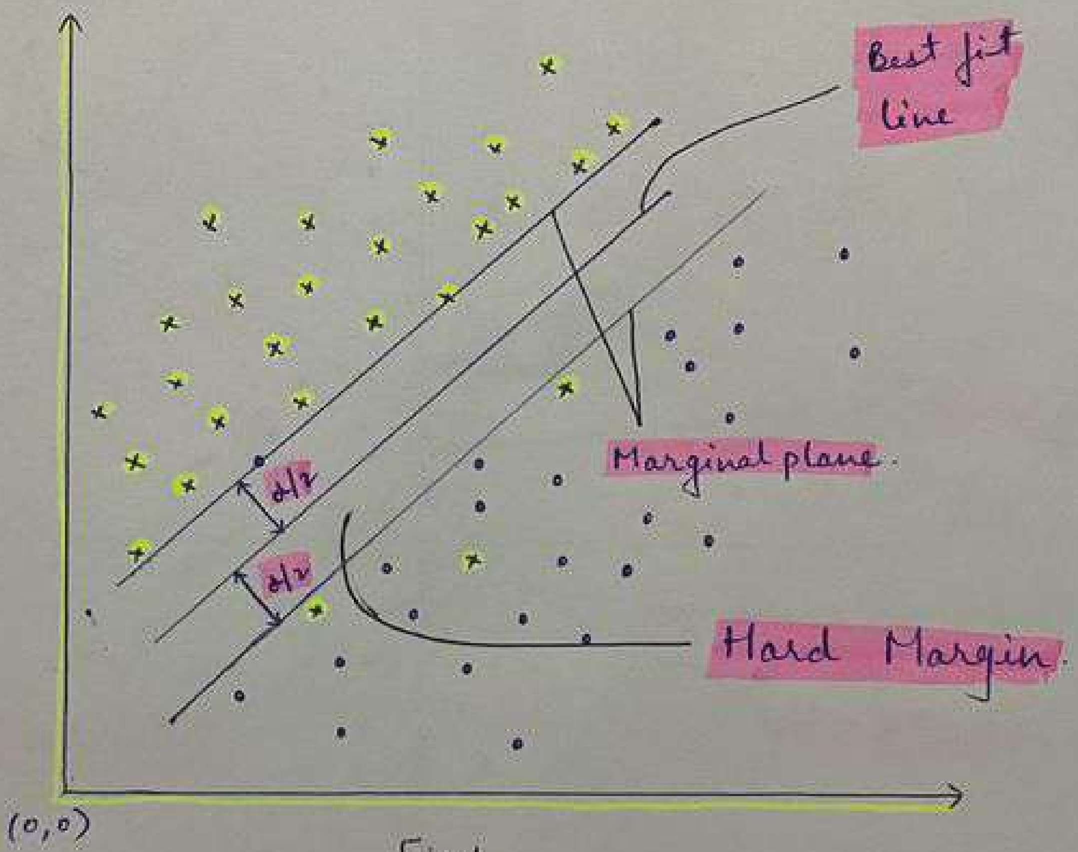


Fig. 1

Marginal plane is equidistance from best fit line.

when there is not data points b/w the marginal plane then it's Hard margin, Means No - Error.

When there is data points b/w the marginal plane then it's soft margin, Means there is Error.

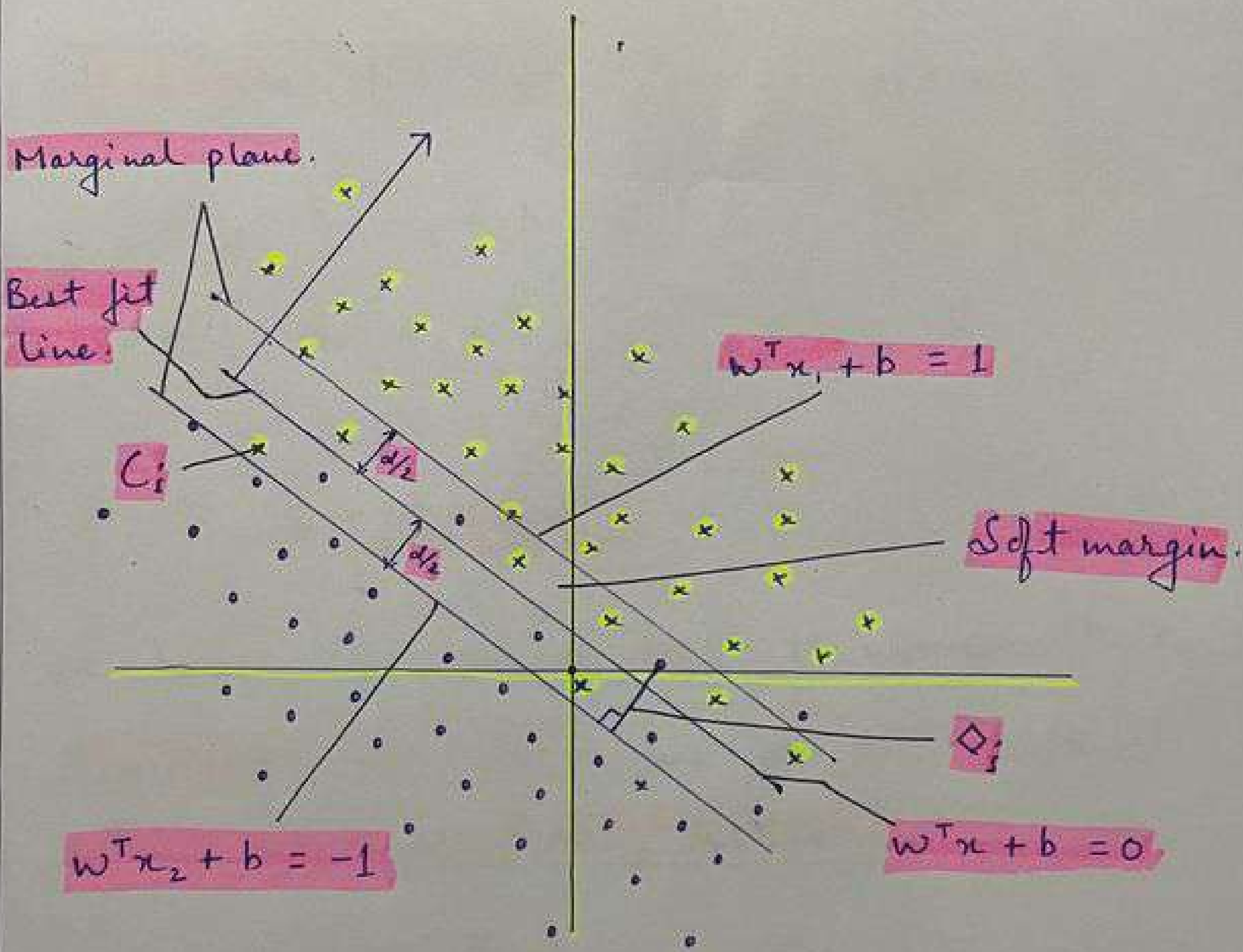


Fig. 2

In real life it's tough to have hard margin or No Error.

let's consider distance in the direction of the $\vec{w}$ plane which is positive and consider unit movement. Consider same in the opposite direction. Movement in both the direction be +1 and -1 (Unit movement).

Maximum distance b/w the planes be:

$$w^T x_1 + b = +1$$

$$w^T x_2 + b = -1$$

$$- \quad - \quad +$$

$$= \quad w^T (x_1 - x_2) = 2$$

we have considered unit movement, so dividing by $|w|$ we get the ~~vector~~ actual marginal plane distance.

$$\boxed{= \frac{w^T (x_1 - x_2)}{|w|} = \frac{2}{|w|}}$$

To get best hard margin we need to maximize the distance $\left(\frac{2}{|w|}\right)$

Constraint:

$$y_i \times \begin{cases} 1 & : w^T x + b \geq 1 \\ -1 & : w^T x + b \leq -1 \end{cases}$$

For datapoints lying in direction of $\vec{w}$

Constraint above marginal plane is +ive

similarly for opposite direction it's -ive.

For all correct points means datapoints and marginal plane both are in same direction of best fit line.

$$\text{Constraint} \Rightarrow y_i \times (w^T x + b) \geq 1$$

In case of cost function we are focused on minimizing. Thus to get max. distance we need to minimize

$$\boxed{\underset{(w,b)}{\text{Minimize}} = \frac{|w|}{2}}$$

Cost function : Hindge loss

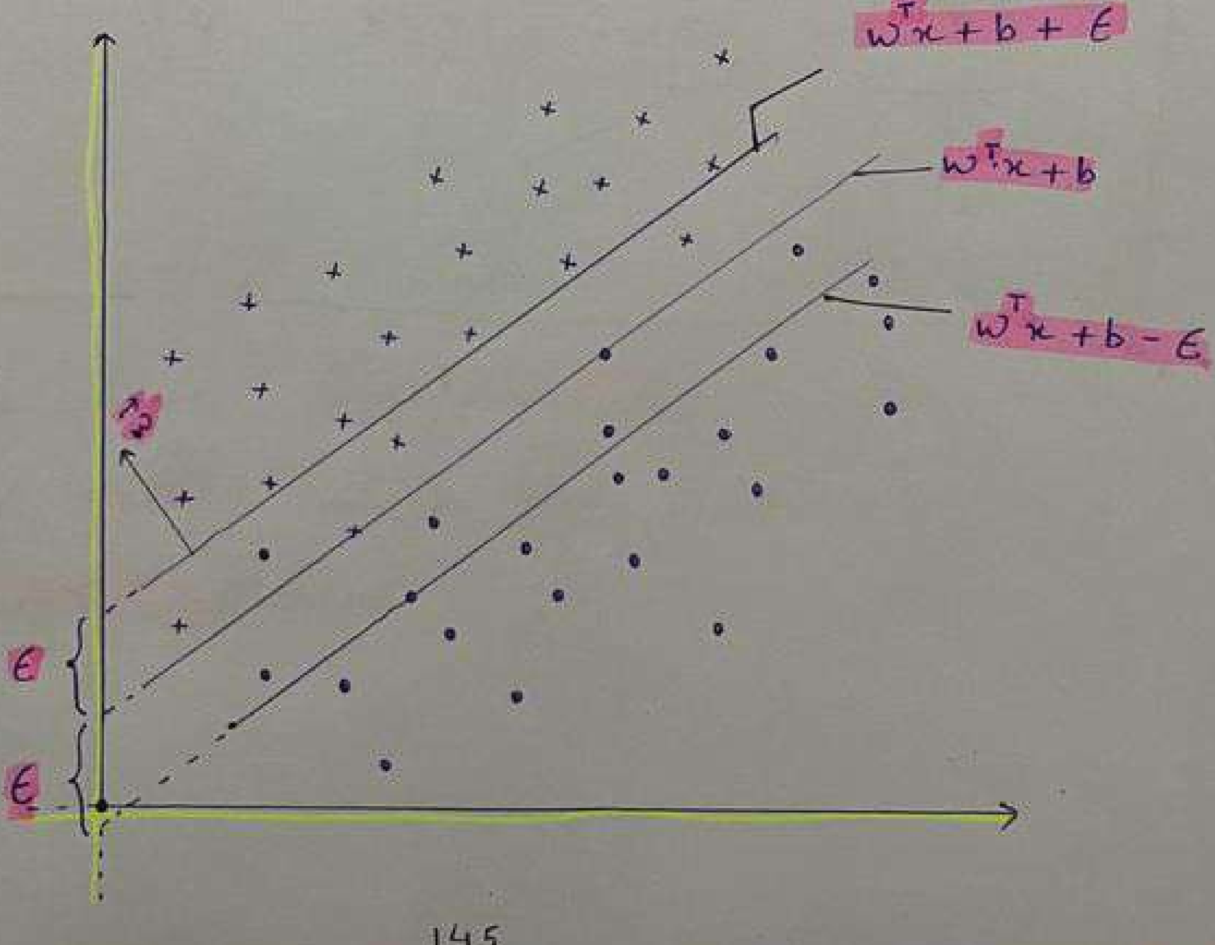
$$\boxed{\underset{(w,b)}{\text{Min}} \quad \frac{|w|}{2} + C_i \sum_{i=1}^{n} \Diamond_i} \quad \text{Eta } i$$

C_i : Sum of datapoints lying b/w the marginal plane.

$\diamond_i$: Distance b/w the datapoints and Marginal plane. Datapoints must lies b/w Marginal plane. Distance is b/w datapoints & Marginal plane of corresponding datapoints.

C_i & $\diamond_i$ is shown in Fig-2

Support Vector Regression.



ϵ : Marginal error. It's error or distance within even if the data Points fall we will consider the output as acceptable.

To minimize the cost function we need to minimize the distance b/w actual and Predicted points.

$$\hat{y}_i = |y_i - w_i x_i|$$

To get best marginal plane:

$$|y_i - w_i x_i| < \underbrace{\epsilon}_{\text{Marginal error}} + \underbrace{|\xi_i|}_{\text{Eta's}}$$

- we have considered ξ_i because these are points which lies outside the marginal plane.

Types of SVM:

- **linear SVM:** Used for linearly seperated data, means dataset can be classified into two classes using a straight line.

- **Non-linear SVM:** Used for non-linear seperated data & data can't be classified using straight line.

Non - linear points:

Consider 1-D points on the X-axis as below:

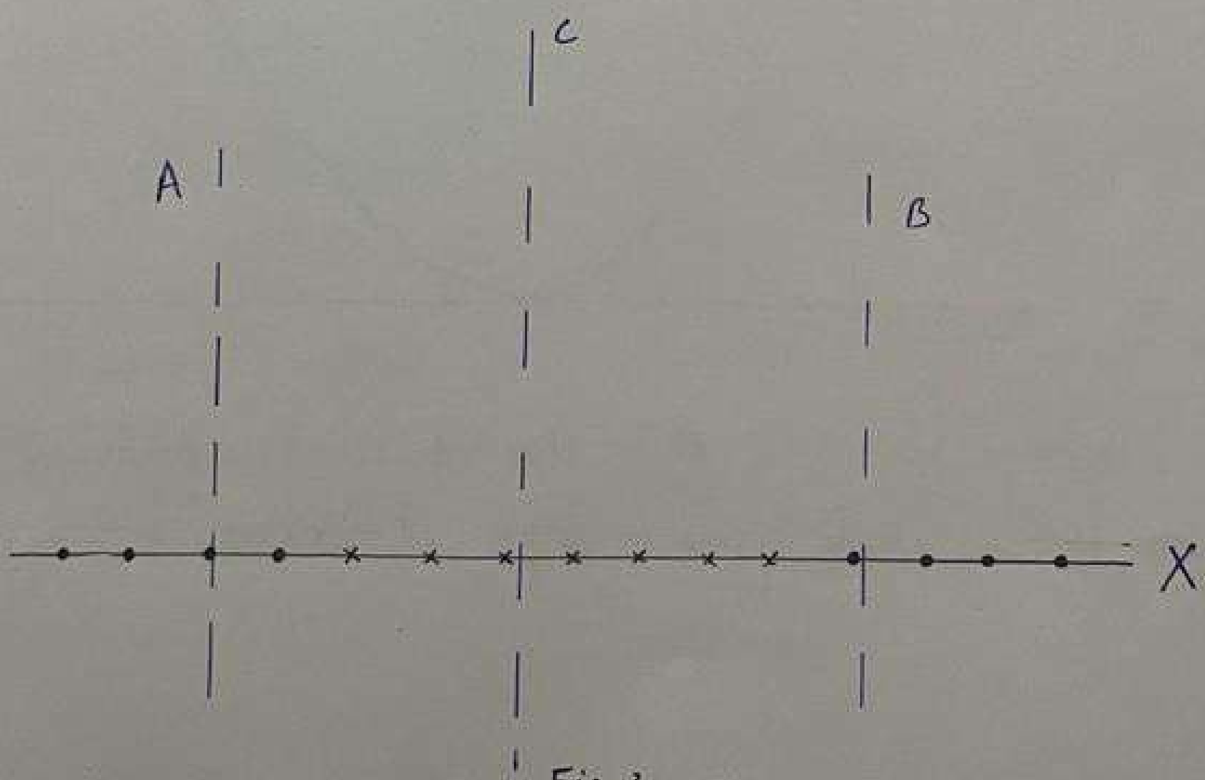
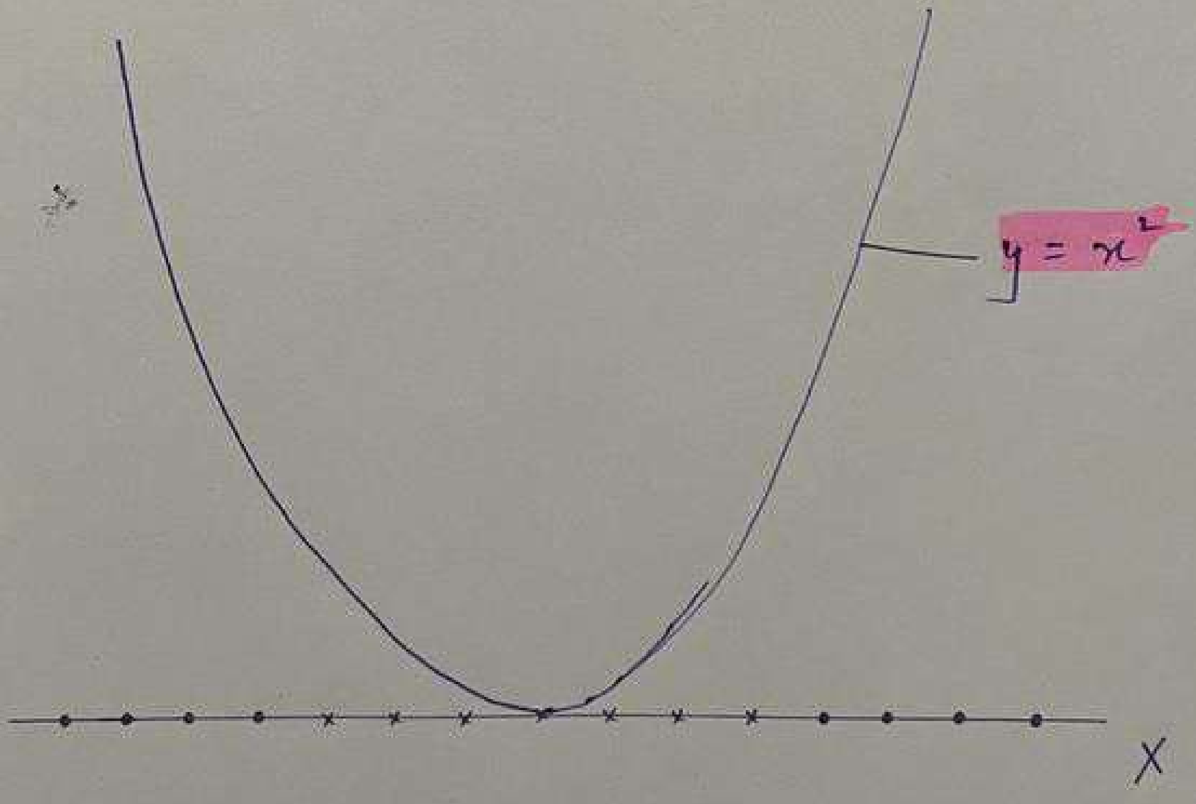


Fig. 3

For above types of data points, to find the best fit line we need to split data points in A and B by the line C. But it's not possible to have two best fit line. So we will transform 1-D to 2-D.



Converting 1-D points to 2-D we make it predictable to for the data points.